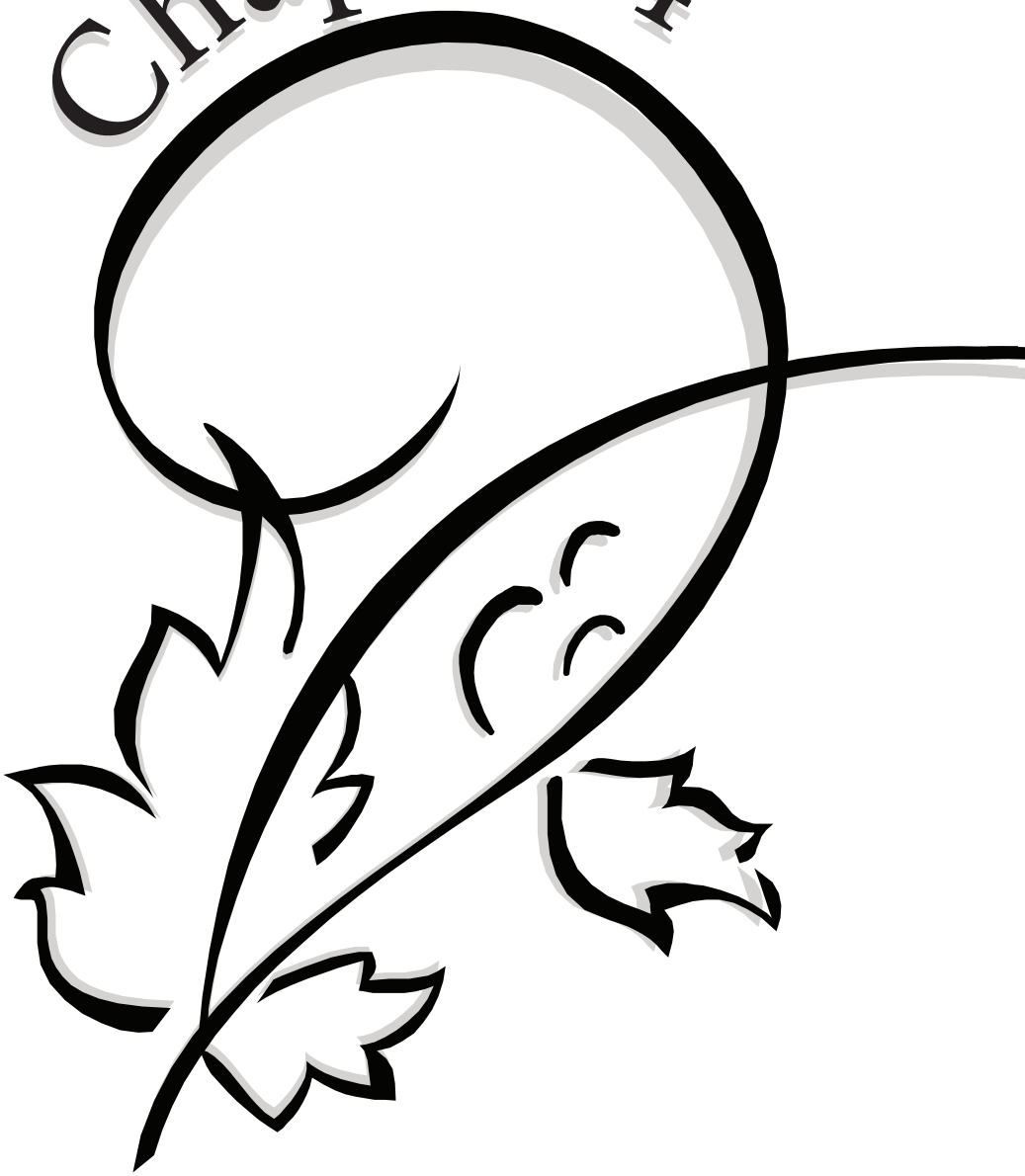
Secondary metabolism by industrially improved Penicillium chrysogenum strains

Salo, Oleksandr

Chapter 4



Identification of a polyketide synthase involved in sorbicillin biosynthesis by *Penicillium chrysogenum*

Oleksandr Salo[1,2], Fernando Guzmán-Chávez[1,2], Marco I. Ries[4,5], Peter P. Lankhorst[6], Roel A.L. Bovenberg[2,6], Rob J. Vreeken[4,5], Arnold J.M. Driessen[1,3,*]

[1]Molecular Microbiology and [2]Synthetic Biology and Cell Engineering, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, The Netherlands

[3]Kluyver Centre for Genomics of Industrial Fermentations, Julianalaan 67, 2628BC Delft, The Netherlands

[4]Division of Analytical Biosciences, Leiden/Amsterdam Center for Drug Research, Leiden, The Netherlands

[5]Netherlands Metabolomics Centre, Leiden University, Leiden, The Netherlands

[6]DSM Biotechnology Center, Alexander Fleminglaan 1,2613 AX Delft, The Netherlands

Abstract

Secondary metabolism in *Penicillium chrysogenum* has been intensively subjected to the classical strain improvement program (CSI) that resulted in the development of industrial strains producing high levels of β-lactams. During this process, yellow pigment production including sorbicillinoids was eliminated as part of a strategy to enable the rapid purification of β-lactams. Here we report the identification of the polyketide synthase (PKS) genes essential for sorbicillinoids biosynthesis in *P. chrysogenum*. We demonstrate that the production of polyketide precursors like sorbicillinol and dihydrosorbicillinol as well as their derivatives bisorbicillinoids require the function of a highly reducing PKS encoded by the gene *Pc21g05080 (pks13).* This gene belongs to the cluster that has been mutated and transcriptionally silenced during the strain improvement program. In the background of an improved β-lactam producing strain, repair of the mutation in *pks13* led to the restoration of sorbicillinoids production. This now enable genetic studies on the mechanism of sorbicillinoid biosynthesis in *P. chrysogenum* and opens new perspectives for pathway engineering.

Introduction

Sorbicillinoids are the diverse group of yellow pigments produced by *Trichoderma* [1]*, Aspergillus* [2], *Verticillium* [3], *Streptomyces* [4] and *Penicillium* [5] species. Sorbicillin (Fig. 1, compound 1) was the first characterized sorbicillinoid initially isolated from *P. notatum* as a contaminant during the production of clinical penicillins [5]. The typical hexaketide structure of this molecule is a core scaffold for more than 30 monomeric and dimeric derivatives isolated from different environments [6]. The oxidative dimerization of the sorbicillinol (2), a hydroxylated derivative of sorbicillin [7], can be achieved via Diels-Alder or Michael type oxidative dimerization reactions [8] leading to structural diverse bioactive compounds with promising bioactive properties. For instance, radical scavenging properties have been assigned to bisorbicillinoids like bisorbicillinol (5), bisvertinoquinole (9) and bisorbibutenolide (10) [9]. The group of trichodimerols (11) shows anti-viral and anti-inflammatory activity by inhibiting the prostaglandin H synthase-2 and tumor necrosis factor (TFN-α) in human peripheral blood monocytes [8]. Bisvertinols (12) are equipped with antimicrobial activity through inhibition of 1,6-glucan biosynthesis in the plant pathogen *Phytophthora capsici* [10]. Recently*,* a sponge-associated *P. chrysogenum E01-10/3* strain was isolated, showing under optimized cultivation conditions, the production of large quantities of sorbicilactone A/B (13,14). These have anti-HIV properties and show cytotoxic effect against L5178y leukemic cells [11, 12]. The remarkable bioactive potential of sorbicillinoids have raised interest in their synthetic origin [13], as well as biosynthetic mechanism. Feeding experiments with radiolabeled acetate indicates that sorbicillinol and dihydrosorbisillinol act as precursors for the corresponding sorbicilactones. For the biosynthesis of these hexaketide precursors, the involvement of a highly reducing (HR) and a non-reducing (NR) PKS enzyme has been proposed, and a putative biosynthetic gene cluster has been suggested for *Penicillium E01-10/3* (Fig. 2a) [14, 15]. An analogous gene cluster exists in *P. chrysogenum*, and consists of seven genes encoding two fungal specific transcriptional factors (Orf1 and Orf5), an oxidase (Orf7), two oppositely transcribed NR and HR PKS enzymes (SorB and SorA), a transporter (Orf6) and a monooxygenase (SorC). The latter enzyme of *Penicillium E01-10/3* has recently been expressed in *E. coli* and shown to catalyze the hydroxylation of the sorbicillin and dihydrosorbicillin yielding sorbicillinol and dihydrosorbicillinol, respectively [7]. The hypothetical biosynthesis pathway proposed suggests that a triacetic product

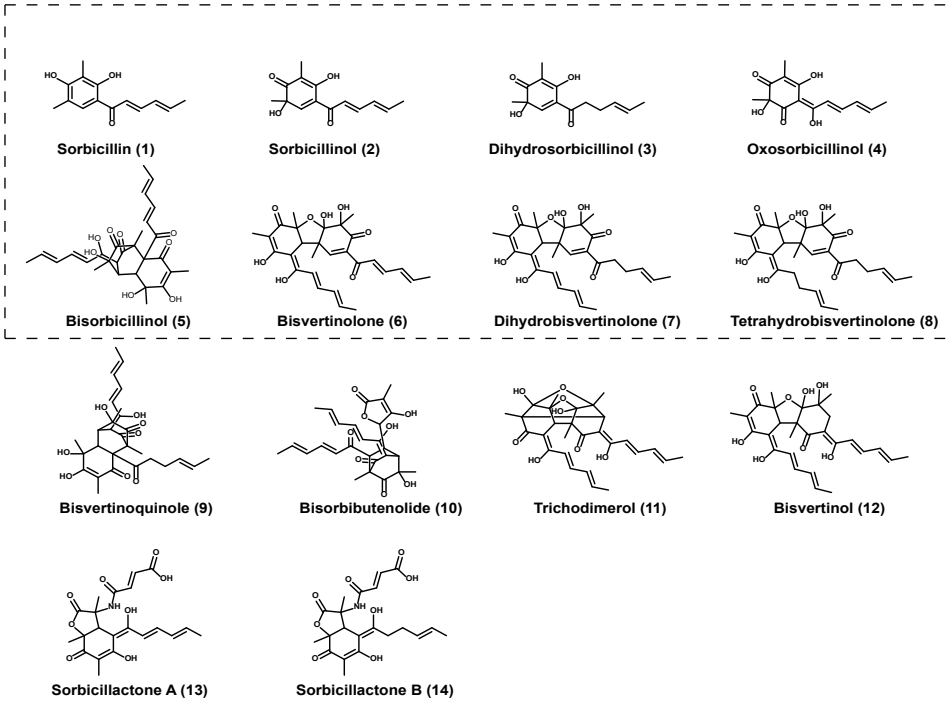


Figure 1. Sorbicillin related compounds isolated from *Penicillium* species. The compounds detected in this study are shown framed.

of the HR-PKS serves as the substrate for the starter unit acyltransferase (SAT) domain of the second NR-PKS enzyme. Upon three iterative malonyl-CoA extensions, the methylated hexaketide might be reductively released from the PKS as an aldehyde and upon cyclization, sorbicillin and/or dihydrosorbicillin are formed. The latter intermediate is presumably derived from a triketide precursor wherein the first enoyl reduction during chain extension by the HR-PKS is omitted (Fig. 2b). Although these studies provide a first glimpse on the possible mechanism of sorbicillinoids biosynthesis, the direct involvement of the PKS enzymes has not been demonstrated.

Unlike natural isolates of *P. chrysogenum,* strains with an improved penicillin production like *Wisconsin 54-1255* and its derivatives are not capable of sorbicillinoids production. Transcriptional profiling of secondary metabolite genes in related strains of a lineage of improved β-lactam producer, indicated the presence of a homologous PKS gene cluster that has been silenced in *Wisconsin 54-1255* during

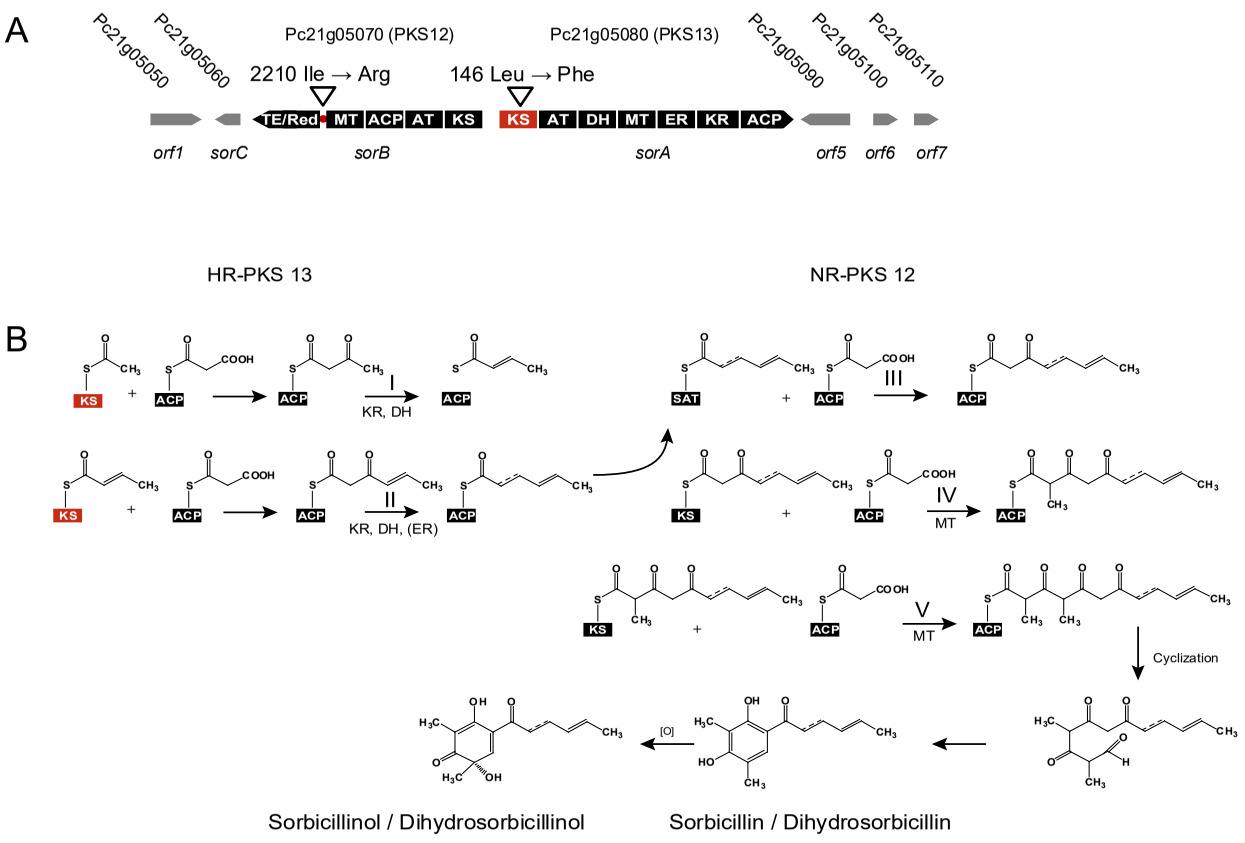


Figure 2. A) Proposed gene cluster of sorbicillinoids biosynthesis in *P. chrysogenum.* Abbreviations used for PKS domains: SAT, starter unit: acyl-carrier protein transacylase (SAT) domain; KS, ketosynthase; AT, acetyltransferase; ACP, acyl carrier protein; DH, dehydratase; KR, ketoreductase; ER, enoylreductase; MT, methyltransferase; TE/Red, thioester reductase domain. **B) Proposed mechanism of sorbicillin/ dehydrosorbicillin biosynthesis (adopted from [14] involving PKS12 and PKS13.**

classical strain improvement (CSI) but displaying the highest transcriptional level in progenitor *NRRL1951* strain that vigorously produces sorbicillinoids (Salo *et al*, submitted). In addition, mutations emerged in each of the putative PKS enzymes that likely led to an inactivation. These findings suggest a complex mechanism for the elimination of biosynthetic pathway of sorbicillinoids during the strain improvement program. Here we report on the identification of the PKS encoding gene *pks13* that is essential for sorbicillinoids biosynthesis using the natural producer strain *P. chrysogenum NRRL1951.* Repair of the native nucleotide sequence of this gene in a strain that was derived from a high β-lactam producing strain resulted in the restoration of sorbicillinoids biosynthesis. This now allows the study of sorbicillinoids biosynthesis using standard molecular techniques that was previously restricted due to the natural genetic background of sorbicillinoids producing isolates.

Materials and Methods

Strains, media and growth condition

The parental *P. chrysogenum* isolate *NRRL1951* and its derivative penicillin gene cluster free strain DS68530 were kindly provided by DSM anti-infectives (Delft, The Netherlands). YGG medium was used to grow the fungus for gDNA extraction and protoplasting. The secondary metabolite production medium (SMP) was used for secondary metabolites analysis. All cultivations were performed in shaken flasks at 25°C and 200 rpm. Positive selection of the transformants was performed on AMDS agar medium supplemented with acetamide, as the nitrogen source. The solid SMP medium supplemented with agar has been used for the rapid selection of the pigment producing fungal colonies. R-agar sporulating medium was used for purification of the transformants and preparing of the rice batches for the long-term storage of the conidia [16].

Construction of the Pc21g05080 gene inactivation strain

A deletion plasmid for the *Pc21g05080* gene was constructed using the modified Gateway™ cloning protocol (Invitrogen). 3'- and 5'- homologous regions for the deletion cassette were amplified with the PCR Master Mix™ with the primers listed in Table 1. The obtained fragments were cloned into the corresponding do-

nor vectors pDONR P4-P1R and pDONR P2R-P3 BP clonase II™ enzyme mix (Invitrogen). The resulting plasmids were isolated from kanamycin resistant *E. coli* DH5α transformants. The constructs were further used for an in-vitro recombination reaction with the destination vector pDEST-amdS (LR clonase II™ enzyme mix) resulting in the isolation of the final pKO13 deletion plasmid that was derived from ampicillin resistant *E. coli* transformants. Before transformation into protoplasts of *P. chrysogenum NRRL1951*, the deletion cassette has been linearized with *AjiI* and *SapI* restriction nucleases (Thermo Scientific). For the targeted integration of the *amdS* marker into the locus of the pks17 (*Pc21g16000*) gene the plasmid pKO17 (Salo *et al*., unpublished) has been used as the template for the amplification of the cassette. The corresponding pair of primers is listed in Table 1.

Transformation and screening

Protoplastation and transformation of the fungal mycelia was done as described previously [17]. Transformants were obtained by inoculation of single conidia on AMDS selective medium followed by a sporulation step on R-agar plates. After two sporulating phases, fungi were grown on rice for inoculation and long-term storage of the conidia. The correct knockout strain was selected by colony PCR analysis using Phire® Hot Start II DNA Polymerase (Thermo Scientific, USA). The mycelium of the transformants was homogenized in 20 μL of milliQ water and 2 μL of the cell suspension was used immediately for PCR validation with the forward primer (pks13F) that anneals outside the recombination region while the reverse primer (amdsR) was *amdS* marker specific (Table 1). The presence of an expected 1.6 kb PCR product was used to select the correct homologous recombinants.

Southern analysis

The downstream region of the *Pc21g05080* gene was used as a probe and amplified by PCR with primer set listed in Table 1. The probe was lled with digoxigenin using the HighPrime™ Kit (Roche Applied Sciences, The Netherlands). gDNA (10 μg) was digested with appropriate restriction enzymes and separated on 0.8% agarose gel. After equilibration in 20x SSC buffer (3M sodium chloride; 0.3M sodium citrate), the DNA was transferred overnight onto Zeta-probe™ positively charged nylon membranes (BioRad). Blots were treated with anti-DIG-alkaline phosphatase

antibodies (Sigma) and supplemented with CDP-star™ (Roche Applied Science, The Netherlands). The fluorescence signal was measured with a Lumi Imager™ (Roche Applied Science, The Netherlands).

Restoration of pks12/13 genes

To repair the mutations in the *Pc21g05070* and *Pc21g05080* genes in strain *DS68530*, a DNA fragment of 9,769 bp covering the two point mutations in these genes was amplified by PCR from genomic DNA from strain *NRRL1951* using the primer set listed in table 1. The DNA fragment was cloned in the vector pJET1.2/blunt (Thermo Scientific™ CloneJET™ PCR Cloning Kit) according to the manufacturer's instructions. The obtained plasmid was purified from ampicillin resistant E. coli DH5α, and used as a new DNA template with the primers above described before. The amplified DNA fragment was used along with the amdS selection marker in a cotransformation of protoplast isolated from P. chrysogenum *DS68530* using a standard protocol [17]. The amdS gene was under control gpdA promoter of A. nidulans. This cassette marker was amplified from plasmid pKO17 by PCR using the primer set listed in table 1, yielding a 1000 bp overhang size in the two flanks of the fragment to target the polyketide synthase gene *Pc21g16000* that is responsible for green pigment formation. The screening for transformants was performed on plates supplemented with acetamide as unique nitrogen source among the mutants had lost green pigmentation. An individual colonies were gown on Solid SMP medium and examined for the formation of a yellow halo. Positive colonies were grown in 10 mL SMP using shaken flasks at 25°C and 200 rpm. After 5 days of shaking, the flasks were supplemented with 8 mL with fresh SMP and growth was continued for 4 days whereupon yellow pigment formation was verified visually and by LC-MS as described below. Restoration of the mutations in *Pc21g05080* and *Pc21g0570* were confirmed by sequencing of the genomic DNA of the clones, using gDNA from strain *DS68530* as a control.

Metabolite analysis

The spent medium of fungal cultures on SMP media was collected after 3, 5 and 7 days growth and subjected to secondary metabolite analysis. Samples were filtered with a 2 µm PTFE syringe filter and stored at -80°C conditions. LC/MS anal-

ysis was performed using Accella1250™ HPLC system coupled with the benchtop ES-MS Orbitrap Exactive™ (Thermo Fisher Scientific, San Jose, CA). A sample of 5 µL was injected into Shim-pack XR-ODS™ C18 column (3.0 x 75 mm, 2.2 µm) (Shimadzu, Japan) operating at 40°C and flow 300 µL/min. The linear gradient began with 90 % of solvent A (100% water) and 5 % of solvent C (100% Acetonitrile) starting after 5 minutes of isocratic flow. The first linear gradient reached 60 % of C at 30 minutes, the second - 95 % of C at 35 minutes. The washing step for 10 minutes at 90% of solvent C was followed by the column equilibration for 15 minutes at initial isocratic conditions. Solvent D (2% Formic acid) was continuously used to maintain the final 0.1 % of formic acid in the system. The column fluent was directed to the Exactive™ ES-MS Orbitrap operating at the scan range (m/z 80 – 1600 da) and switching positive / negative modes. Voltage parameters for positive mode were: 4.2 kV spray, 87.5 V capillary and 120V of tube lens. Voltage parameters for negative mode were: 3kV spray, -50 V capillary, -150 V tube lens. The capillary temperature of 325°C and a sheath gas flow of 60 a.u. was used. Auxiliary gas was off to maintain a high detection sensitivity for both positive and negative modes during analysis. The differential analysis of the LC-MS samples was performed using the Thermo Scientific™ SIEVE software.

NMR

NMR spectra were recorded on a Bruker Avance III 700 MHz or 600 MHz spectrometer. Sample temperature ranged from 250 – 300K. The assignments were achieved by means of 1D ^{13}C, COSY, TOCSY, HSQC, and HMBC spectra. Samples were dissolved in $CDCl_3$, and in $CDCl_3$ with one drop of pyridine to neutralize acidic impurities in the chloroform.

Results

Deletion of the *pks13* in *P. chrysogenum NRRL1951*

P. chrysogenum NRRL1951 is a low β-lactam producing strain that has not been subjected to extensive classical strain improvement (CSI) and that secretes yellow pigment compounds into the medium. These yellow pigments originate from the hexaketide sorbicillin, but the polyketide synthase responsible for its production

has not been identified. Previously we demonstrated that corresponding metabolites are produced early during fermentation which is accompanied with the expression of a PKS gene cluster of unknown function (Salo et al., submitted). Protein BLAST analysis indicated that particular cluster consists of the two PKS genes (*Pc21g05070*, *pks12*; *Pc21g05080*, *pks13),* two transcription factors (***Pc21g05050*** and ***Pc21g05090***, which are termed *reg50* and *reg90,* respectively), a monooxygenase (***Pc21g05060***, *mox60*), a transporter (***Pc21g05100***, *mfs100*), and an oxidoreductase (***Pc21g05110***, *ox110*) (Fig. 2a) The predicted product of *pks12* is a 2581 amino acid long non-reducing iterative type I polyketide synthase showing the highest (64 %) identity to an unknown PKS gene of *Trichoderma reesei* and 43 % identity to citrinin biosynthesis gene of *Monascus purpureus* [18]. The neighboring *pks13* gene encodes a 2664 amino acid long highly reducing polyketide synthase with 65 % identity to an unknown PKS of *T. reesei* and 48% identity to the lovastatin diketide synthase LovF of *A. terreus* [18]. In the derived strains *P. chrysogenum Wisconsin 54-1255* and *DS17690*, the two PKS genes within the aforementioned cluster acquired mutations during the CSI that may lead to their functional inactivation. In the non-reducing PKS Pc21g05070 (here and after termed *pks12*) the isoleucine at position 2210 located at inter-domain region of methyltransferase (MT) and thioesterase (TE) domains is substituted for an arginine. The second mutation is present within the ketosynthase domain of highly reducing PKS Pc21g05080 (here and after termed *pks13*) causing a leucine to phenylalanine substitution at position 146 (Fig. 2ac). Both mutations occurred early during the CSI and have been inherited by the *Wisconsin 54-1255* strain and thus also in later improved β-lactam producers (Salo et al., submitted) (Fig. 2a). Considering that the CSI derived strains of *P. chrysogenum* are not able to produce yellow pigments, the natural isolate *NRRL1951* was chosen as the host in our study. To functionally characterize the aforementioned cluster, the PKS encoding gene *pks13* was targeted for inactivation using homologous recombination and by replacing it for the *amdS* marker for acetamide selection. However, gene targeting in the *NRRL1951* strain occurs with a very low efficiency due to the predominance of non-homologous recombination events. Therefore, to identify the desired gene inactivation mutant, among the majority of random insertion mutants, PCR screening was applied. Using this approach, a PCR product of 1.6 kb (See material and method section) indicated the correct homologous recombination event, which could be assigned to one out of 32 transformants that passed two sporulating steps on acetamide selective medium.

Purified gDNA of this strain was used for Southern blot analysis, which confirmed the correct *pks13* gene deletion (Fig. 3).

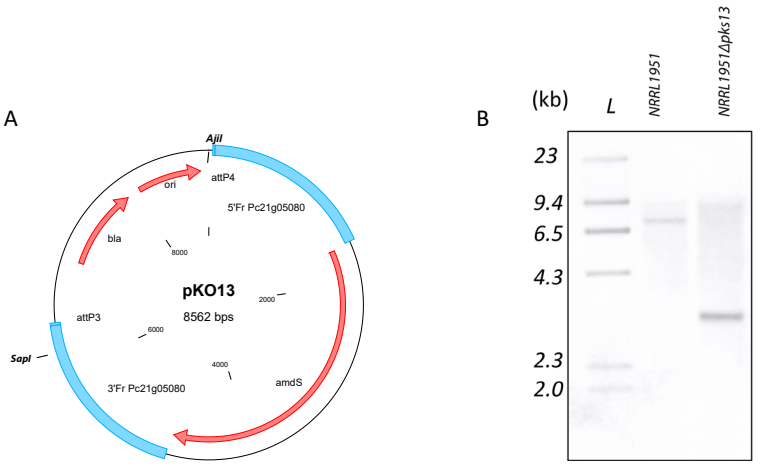


Figure 3. Schematic representation of the *pks13* gene deletion and its confirmation by southern blot analysis. A) Scheme of the deletion plasmid pKO13. Features: *amdS*, an *A. nidulans* acetoamidase gene for positive selection of the fungal transformants on acetamide supplemented medium as a sole nitrogen source; *bla,* ampicillin resistance gene for the selection in *E. coli; ori*, pUC origin of replication; attB3/4, Gateway recombination sites. B) Southern blot analysis. Genomic DNA was digested with NdeI endonuclease, using *pks13* 3' FR as a probe. The expected 7.3 kb DNA fragment of the parental strain *NRRL 1951* and 3.6 kb signal confirming that *pks13* locus is replaced with amdS marker.

Phenotype of the Δpks13 strain and secondary metabolite analysis

To verify the effect of the *pks13* gene deletion on secondary metabolism, the mutant *P. chrysogenum NRLL1951* strain (here after termed *NRRL1951Δpks13*) and the parental strain were grown in liquid SMP medium. The *NRRL1951Δpks13* mutant was not able to produce the typical yellow pigmentation compare to the parental strain that exhibited a yellow coloring of the culture already after three days of growth (Fig. 4a). There were no further phenotypical differences detected between both strains. For secondary metabolite analysis, samples of the culture

broth were obtained after 3, 5 and 7 days of cultivation. The corresponding metabolite profiles of the *NRRL1951* and *NRRL1951Δpks13* strains were analyzed by HPLC-MS. Comparative analysis indicated that a group of metabolites is absent in the culture medium of the *NRRL1951Δpks13* strain. The empirical formula of these compounds were calculated based on the detected accurate mass (ppm<2). The major compound produced after 3 days of growth has a retention time RT 21.15 min and m/z $[H]^+$ 249.11, with the calculated empirical formula $C_{14}H_{16}O_4$ which corresponds to sorbicillinol (2). The second compound with RT 23.47 min and a m/z $[H]^+$ 251.13 has the empirical formula $C_{14}H_{18}O_4$ and corresponds to dihydrosorbicillinol (3). Both masses were found to be part of the MS/MS fragmentation pattern of compounds (4), (6) and (5)**,** respectively, while (2) shows a corresponding fragmentation pattern overlapping with (3). This indicates that (3) is composed of (1) and (2). A complete list of the unique masses related to the *pks13* deletion is shown in Table 2. To confirm that the compounds eliminated from the secondary metabolism of the *NRRL1951Δpks13* mutant indeed belong to the class of sorbicillinoids, metabolite (5) was isolated by means of preparative HPLC and its structure was verified by NMR. The isolated fraction was dissolved in $CDCl_3$, as well as in $CDCl_3$ with one drop of pyridine-d5 with the purpose to neutralize traces of acid. In the presence of acid the metabolite (5) occured in two tautomeric forms and suffered from a slow degradation, which was not the case in the neutralized NMR sample. 2D spectra were recorded from both NMR samples, and careful interpretation led to the conclusion that metabolite (5) must be bisorbicillinol. All ^1H and ^{13}C NMR chemical shifts agree very well with the data reported previously [9]. In addition these authors also observed the presence of the tautomeric equilibrium, albeit in a slightly different way. They observed two compounds after derivatization with diazomethane and concluded that this was the result of a tautomeric equilibrium. More details of the assignment and NMR data are given in the Supplementary data. Overall, these results indicate that *pks13* is essential for the production of the polyketide precursors for the biosynthesis of sorbicillinoids and their derivatives by the *P. chrysogenum* strain *NRRL1951*.

Recovery of sorbicillinoids biosynthesis in the improved β-lactam producing strain DS68530

In order to define the functional role of the mutations accumulated by the improved penicillin producing strains during the CSI, we restored the native amino acid sequence of the PKS enzymes mutated during the CSI in the industrially improved strain *DS58630.* Strain *DS58630* is a derivative of strain *DS17690* in which the multiple β-lactam gene clusters have been removed genetically [19]. This ensures a secondary metabolite pattern that is not further dominated by the presence of β-lactams. A DNA fragment of 9.7 kb region of the oppositely oriented *pks12* and *pks13* genes was amplified from the genomic DNA of the parental strain *NRRL1951.* This allowed the recovery of the native nucleotide sequence of the mutated *pks13* by a homologous recombination event. For positive selection of the

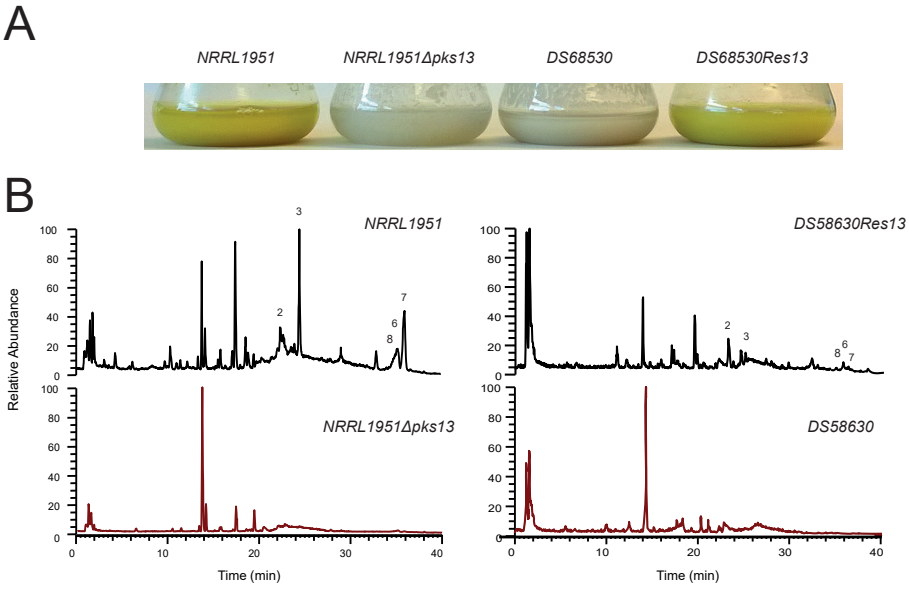


Figure 4. Secondary metabolite profiling of liquid cultures of *NRRL1951* and the deletion strain *NRRL1951Δpks13,* the yellow pigment-less strain *DS58630 and DS58630Res13* carrying the restored native nucleotide sequences of the *pks13* gene. A) Cultures were grown for 3 days on liquid SMP medium in shaking flasks. B) HPLC-MS elution profiles. The major compounds eliminated from the secondary metabolism of *NRRL1951Δpks13* and restored in *DS58630Res13* strain are indicated: (2) sorbicillinol; (3) dihydroxysorbicillinol; (6) bisvertinolon; (7) dihydrobisvertinolone, and (8) tetrahydrobisvertonolone.

Table 1. Primers used in this study.

	Target	Primer sequence (5'- 3')
1	attB4FPc21g05080	**GGGGACAACTTTGTATAGAAAAGTTG**CGTCGGCCGTATTGCCAGACTGC
2	attB1RPc21g05080	**GGGGACTGCTTTTTTGTACAAACTTG**GCCGCTGTTTCACCCGAGTAACC
3	attB2F Pc21g05080	**GGGGACAGCTTTCTTGTACAAAGTGG**GGTCATGTCCGAGAAGCTGTC
4	attB3RPc21g05080	**GGGGACAACTTTGTATAATAAAGTTG**CGCCCTTGTTGAAAGGCTCC
5	pks13F	GGCCGCCATGACAGACTCAGAC
6	amdsR	CACCGGTCACTGTACAGAGCTCG
7	Probe 13 F	GGTCATGTCCGAGAAGCTGTC
8	Probe 13 R	CGCCCTTGTTGAAAGGCTCC
9	pks17F	AATGATACCTTTAGATCTACATTTCCTCACC
10	pks17R	ATTTGGCCGCCGAGAATGAGAGACT
11	Fw-NRRLpks12	GCACTGTCGATATTCAGATGT
12	RV-NRRLpks13	CTTGGTTGAGCATCGATTC

transformants, the AMDS selection marker was co-transformed and targeted to the open reading frame of the naphthapyrone synthase (*Pc21g16000*) that is essential for the green conidial pigment biosynthesis in *P. chrysogenum* (Salo et al., unpublished data). The albino phenotype of the *amdS* carrying mutants provided an additional control over the purity of the transformants grown on the media without selective pressure. Mutants able to grow on acetamide supplemented medium and deficient in green coloring of the conidia were selected after three rounds of sporulation on R-agar and selection on AMDS media. The mycelium of the obtained candidates was inoculated on SMP-agar medium to perform a qualitative selection of the yellow pigment-producing clones. The 15 kb genome region carrying both clustered *pks12* and *pks13* was amplified by PCR and sequenced, and this confirmed the correct reversion of the mutation in PKS13. The corresponding *DS68530Res13* strain was used for the secondary metabolite production analysis. LC-MS analysis of SMP medium grown cultures revealed the accumulation of novel metabolites in the culture medium of strain *DS68530Res13* (Fig. 4b). These were sorbicillin (1) sorbicillinol (2), dihydrosorbicillinol (3), oxosorbicillinol (4), bisorbicillinol (5), bisvertinolone (6), tetrahydrobisvertinolone (7) and dihydrobisvertinolone (8) (Table 2). In addition, a set of new compounds (15-22) that were previously found in the cul-

Table 2. Comparative metabolite profiling for sorbicillin-related compounds in the culture broth of strains *NRRL1951, NRRL1951Δpks13, DS68530* and *DS68530Res13.* The retention time on LC-MS and the calculated empirical formula are indicated.

	Name	Formula	Acquired $[M+H]^+$	RT, (min)	*NRRL1951*	*NRRL1951 Δpks13*	*DS58630*	*DS58630 Res13*
1	Sorbicillin	C14H16O3	233.12	30.80	0.88	-	-	0.08
2	Sorbicillinol	C14H16O4	249.11	21.15	13.93	-	-	1.56
3	Dihydrosorbicillinol	C14H17O4	251.13	23.47	26.85	-	-	1.02
4	Oxosorbicillinol	C14H16O5	265.11	26.90	0.32	-	-	0.15
5	Bisorbicillinol	C28H32O8	497.22	29.43	0.25	-	-	0.02
6	Bisvertinolon	C28H32O9	513.21	32.81	4.81	-	-	0.09
7	Dihydrobisvertinolone	C28H34O9	515.23	35.62	12.18	-	-	0.04
8	Tetrahydrobisvertinolone	C18H36O9	517.24	32.44	10.53	-	-	0.03
15	Unknown	C12H14O3	207.10	23.25	2.91	-	-	0.02
16	Unknown	C11H12O3	193.09	20.62	8.13	-	-	-
17	Unknown	C12H17ON	192.14	13.31	1.28	-	-	0.082
18	Unknown	C15H20O4N2	293.15	17.43	15.56	-	-	3.30
19	Unknown	C15H20O5N2	309.14	15.22	2.84	-	-	0.29
20	Unknown	C16H21O3N3	304.17	13.35	2.68	-	-	0.11
21	Unknown	Unknown	657.26	32.88	0.79	-	-	0.78
22	Unknown	Unknown	657.26	34.14	0.22	-	-	0.21

-, not found.

ture broth of *NRRL1951* strain were also observed in the medium of the *DS68530* mutant. The fragmentation pattern shows that they are related to the sorbillinoids but their structures are unknown. These data demonstrate that the reversion of the mutation in *pks13* suffices to restore sorbicillinoids production in a classical strain improvement selected *P. chrysogenum* strain.

Discussion

The yellow pigments sorbicillinoids are a large group of structurally related metabolites produced by many fungi [1, 3, 5, 11, 20-23]. The polyketide origin of these compounds was demonstrated by radiolabeled acetate feeding experiments [11] but the genes involved in the biosynthesis of these secondary metabolites remained unknown. In this work we deleted the *pks13* (*Pc21g05080*) gene located in a highly expressed gene cluster of *P. chrysogenum NRRL1951* to elucidate its role in secondary metabolism of this fungus. *Pks13* belongs to a cluster of seven genes among which a second *pks* gene that is oppositely transcribed, namely *pks12* (*Pc21g05070*) (Fig. 2). A related gene cluster has been implicated in the biosynthesis of sorbicillacton A/B in the marine isolate *E01-10/3* [7, 11], but direct evidence for the involvement of the PKS enzymes was not demonstrated. As shown in our previous work (Salo *et al.*, submitted), this complete gene cluster is highly expressed in the parental strain *NRRL1951* and is transcriptionally silenced in the improved β-lactam producer *Wisconsin 54-1255* and its derivatives. The mechanism of the transcriptional silencing of this gene cluster is unknown but its functional inactivation in the later generation of *P. chrysogenum* strains that were obtained by classical strain improvement can be assigned to mutagenesis events at the early stages of CSI. Each of the PKS encoded genes carry single nucleotide polymorphism that cause amino acid substitution, i.e., at the intra domain region of PKS12 and within the KS domain of the PKS13 respectively. Remarkably, the leucine residue at position 146 that was substituted for a phenylalanine during the CSI, is conserved within the KS domains of the highly homologous PKSs of other sorbicillin producers like *Trichoderma* and *Glomerella graminicola* (Fig. S7). It is important to stress that identical gene clusters are present in the genomes of these organisms but in *Glomerella graminicola* this is restricted to only the two PKS enzymes.

Since the parental strain *NRRL1951* still produces sorbicillinoids, *pks13* was

targeted for gene inactivation. Correspondingly, a PCR screening approach was applied to select for the correct transformants among the majority of the non-homologous integrants that have randomly incorporating the deletion cassette into the genome. As a result, a deletion mutant *NRRL1951Δpks13* was obtained that was no longer able to produce the yellow pigmentation typical for the parental strain *NRRL1951* (Fig. 4a). Comparative HPLC-MS analysis revealed the absence of a group of metabolites in the culture broth of the *NRRL1951Δpks13* strain that can be characterized as sorbicillinol and dihydrosorbocillinol based on the exact mass and calculated empirical formulas. The characteristic fragmentation patterns (data not shown) of the eliminated metabolites indicated the presence of sorbicillinol or dihydrosorbicillinol moieties incorporated into a large number of other derivatives (Table 2). To prove that the eliminated compounds indeed belongs to sorbicillinoids, NMR analysis was performed for one of the extracted molecules which was verified to be bisorbicillinol.

To investigate the functional role of the amino acid substitutions obtained during the CSI in *pks12* and *pks13*, a reverse mutagenesis approach was applied to the CSI derived strain *DS58630*. Remarkably, the homologous recombination approach led to a restoration of the *pks13* mutation only and not of the mutation in *pks13*. Nevertheless, this sufficed to restore sorbicillinoids production. In the culture broth of the restored strain, sorbicillinol (2) and dihydrosorbicillinol (3) as well as further derived sorbicillinoids were readily detected. This includes compounds (1, 4-8) as well as structurally unknown intermediates (15-22) (Table 2) that based on the fragmentation patterns in LC-MS/MS can be classified as sorbicillinoids. It is unclear why the *pks12* gene was not repaired by the homologous recombination. Possibly, the DNA fragment which covered the relevant parts of both *pks12* and *pks13* is processed or fragmented during the transformation leading to the correction of one mutation only. Since we obtained very similar patterns of sorbicillin formation in the repaired *DS68530* strain compared to the original *NRRL1951* strain, we conclude that the mutation in *pks12* was not disruptive. Possibly, the mutation lowers the activity.

In conclusion, our work identified *pks13* (*Pc21g05080*) as a key gene in the biosynthesis of sorbicillinoids and derivatives in *P. chrysogenum*. We restored the production of the sorbicillin-related metabolites in a genetic background of an improved penicillin producer strain that is adapted for growth under industrial conditions and for which a genetic toolbox is available. This will simplify the application

of the molecular cloning techniques for studying of sorbicillinoids biosynthesis and opens new perspectives to engineer this pathway for the production of individual sorbicillinoids.

Acknowledgements

We acknowledge the financial support from the Netherlands Organization for Scientific Research (NWO) via the IBOS (Integration of Biosynthesis and Organic Synthesis) Program of Advanced Chemical Technologies for Sustainability (ACTS), the Kluyver Centre for Genomics and Industrial fermentation, which is part of the Netherlands Genomic initiative/Netherlands Organization for Scientific Research, and the Perspective Genbiotics program subsidized by Stichting toegepaste wetenschappen (STW). FG is supported by the Consejo Nacional de Ciencia y Tecnología (Conacyt) Mexico.

Supplementary data

NMR analysis

The NMR spectra of one of the isolated sorbicillin-related compounds were obtained in chloroform. All assignments were obtained from the spectrum in $CDCl_3$ with a drop of pyridine. The assignments were compared with the two different sets of NMR signals observed in fresh chloroform alone. From the LC/MS data it was clear that the formula of the compound is $C_{28}H_{32}O_8$, corresponding to trichodimerol which is an oxidized dimer of sorbicillin (Fig. S1). However, NMR data do only partially agree with the NMR data on this compound published by Andrade et al. [24]. Therefore, the 2D spectra were studied in more detail. All data match with bisorbicillinol and the NMR data given by Abe et al. [9]. All correlations in the HMBC were verified, and agreed with the proposed structure by Abe et al. [9]. Only two carbon signals were missing in the HMBC spectrum, i.e. C9 and C11. However, these carbon signals could be detected without doubt in the HMBC spectrum in chloroform alone, which gave rise to a mixture of the two tautomers. From this spectrum it was possible to conclude that the tautomeric equilibrium is between the carbon atoms C9, C10 and C11, as indicated in Figures S2 and S3. The 1D ^1H spectra in CDCl3 with pyridine and in CDCl3 are shown in supplemental Figures S4 and S5. The enolic OH appears at 13.5 ppm, and it is clear that the number of peaks is doubled in the solvent without pyridine. A part of the HMBC spectrum in $CDCl_3$ is shown in supplemental Fig. 6. It is clear from Figure S6, that C11 in tautomer 1 and C9 in tautomer 2 have chemical shifts between 160 and 170 ppm, which is indicative of the enol form. On the other hand C9 in tautomer 1 and C11 in tautomer 2 have chemical shifts between 190 and 200 ppm, which indicates the keto form. Further details of the assignment are not shown here, and the NMR chemical shifts are summarized in Table S1.

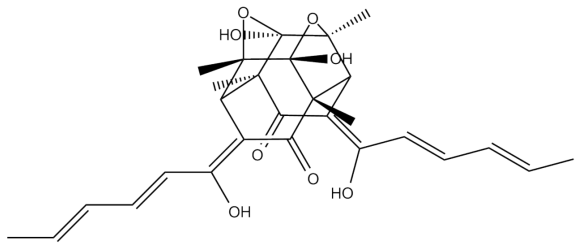


Figure S1. Chemical structure of trichodimerol

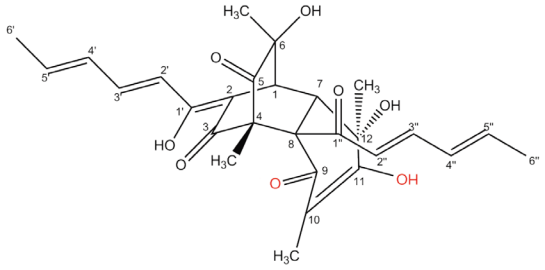


Figure S2. Chemical structure and atom numbering of bisorbicillinol, tautomer 1

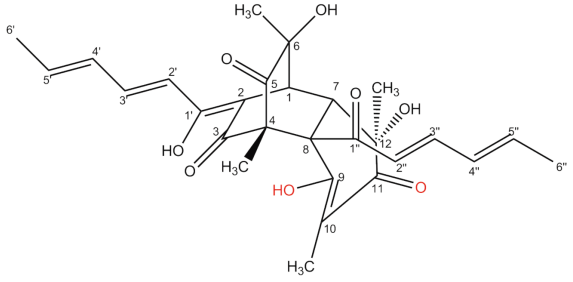


Figure S3. Chemical structure and atom numbering of bisorbicillinol, tautomer 2.

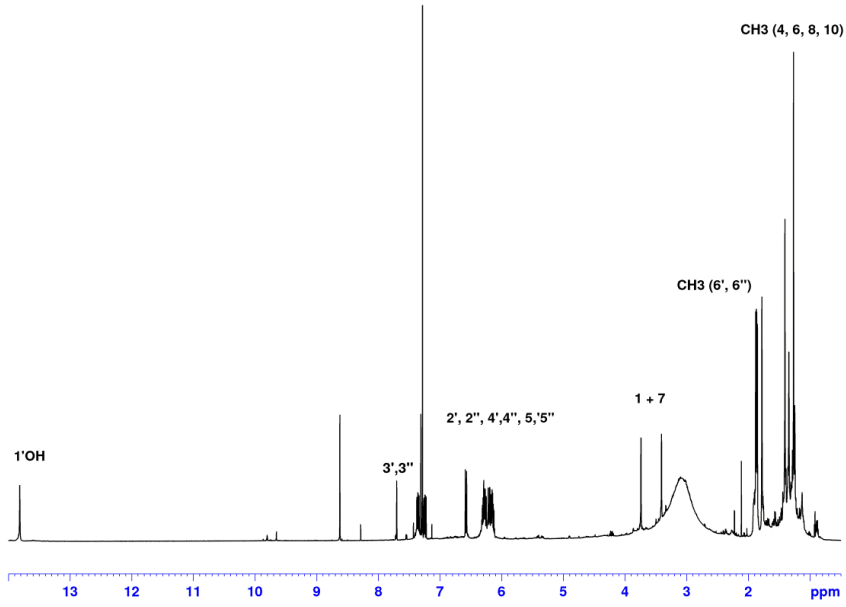


Figure S4. ^{1}H NMR spectrum in $CDCl_3$ with pyridine

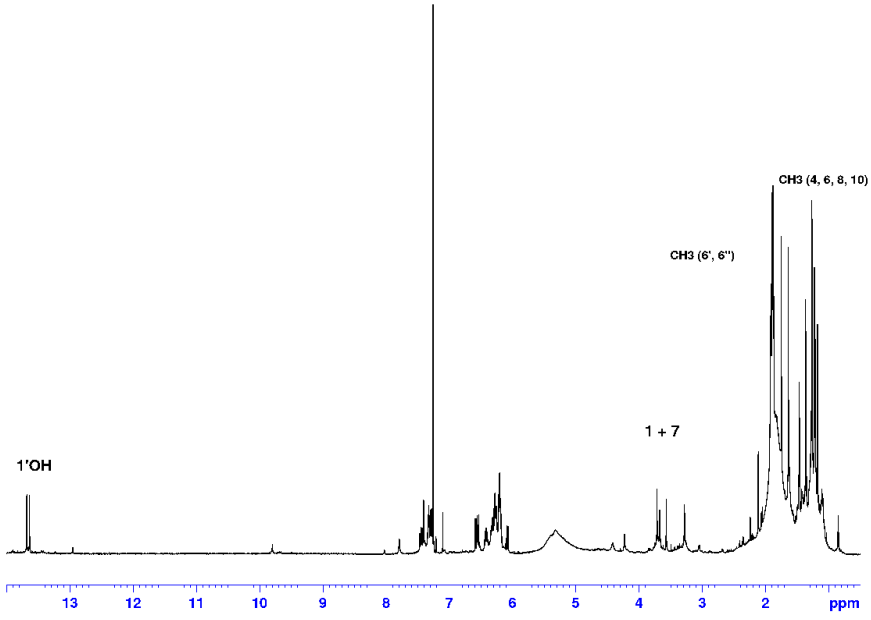


Figure S5. ^{1}H NMR spectrum in $CDCl_3$

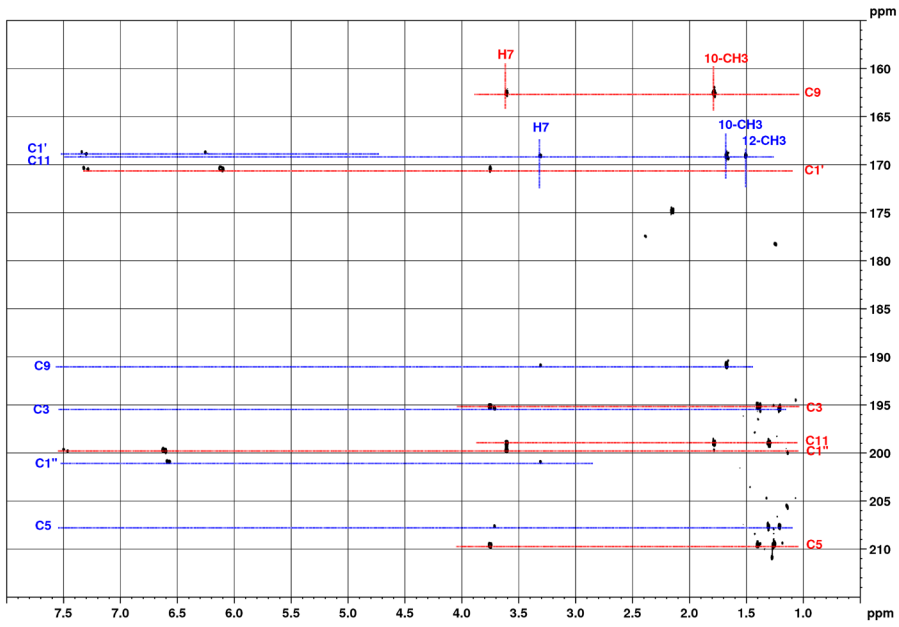


Figure S6. HMBC spectrum (carbonyl signals and enolic signals only) in $CDCl_3$. Blue dashed indicate signals of tautomer 1, red dashed lines indicate tautomer 2

Table S1: Chemical shifts of bisorbicillinol (tautomer 2), 280K, $CDCl_3$ and pyridine-d5. $\delta_{TMS} = 0$

	$^{1}H(\delta)$	$^{13}C(\delta)$
1	3.75	41.4
2	n.a.	108.5
3	n.a.	196.1
4	n.a.	67.9
5	n.a.	208.3
6	n.a.	74.3
7	3.42	47.5
8	n.a.	66.3
9	n.a.	164.5
10	n.a.	110.9
11	n.a.	200.1
12	n.a.	70.5
1’	n.a.	169.2
2’	6.22	118.6
3’	7.24	142.2
4’	6.28	131.1
5’	6.14	139.6
6’	1.90	19.0
1’’	n.a.	200.4
2’’	6.56	123.9
3’’	7.36	146.3
4’’	6.17	130.4
5’’	6.29	143.9
6’’	1.87	19.1
4-CH_3	1.35	10.4
6-CH_3	1.28	24.7
10-CH_3	1.78	9.5
12-CH_3	1.42	33.1

n.a., not applicable

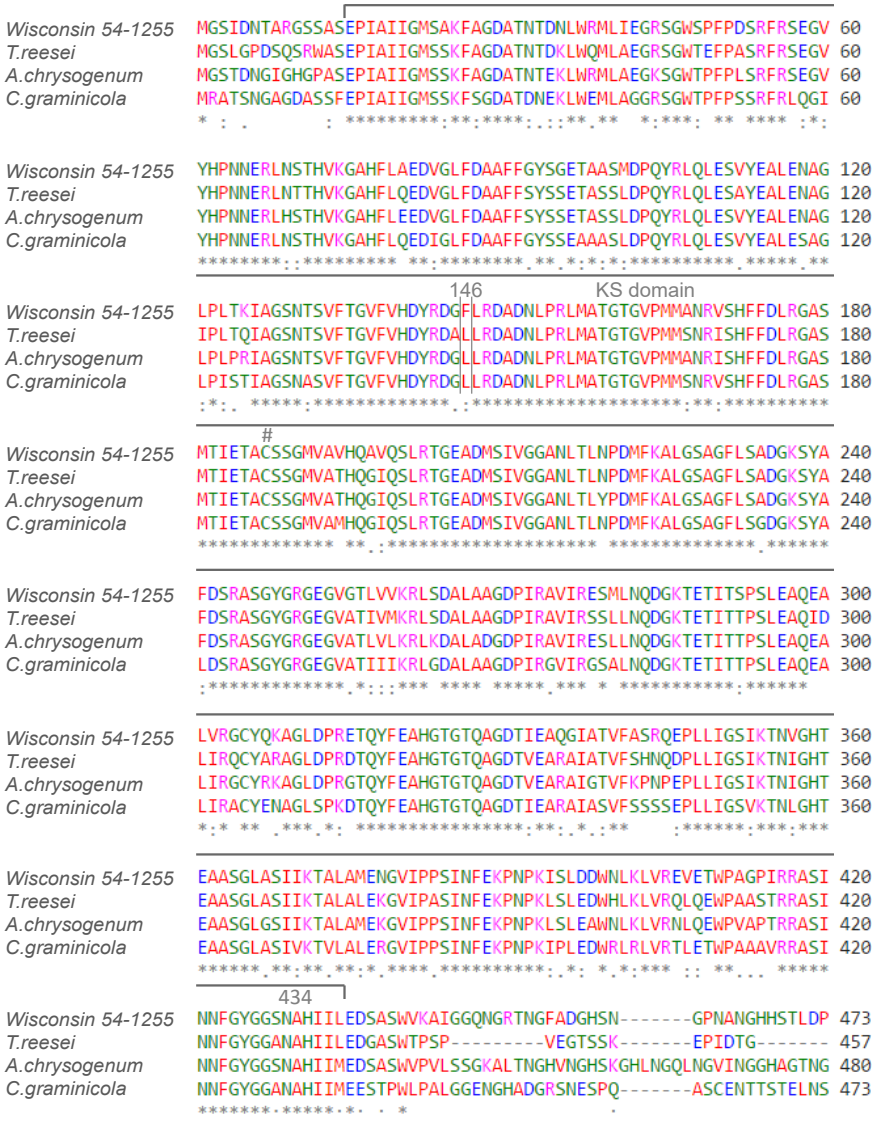


Figure S7. Multiple sequence alignment of the ketosynthase domains of highly homologous PKS enzymes derived from *P. chrysogenum Wisconsin 54-1255, Trichoderma reesei* (66% identity), *Acremonium chrysogenum* (69% identity) and *Colletotrichum graminicola* (63 % identity). The substituted leucine at position 146 is highlighted. The conserved cysteine residue in the active center of the domain is indicated with a hash.

References

1. Abe N, Yamamoto K, Hirota A: **Novel fungal metabolites, demethylsorbicillin and oxosorbicillinol, isolated from Trichoderma sp. USF-2690**. *Bioscience, biotechnology, and biochemistry* 2000, **64**(3):620-622.
2. Basaran P, Demirbas RM: **Spectroscopic detection of pharmaceutical compounds from an aflatoxigenic strain of Aspergillus parasiticus**. *Microbiological research* 2010, **165**(6):516-522.
3. Reategui RF, Wicklow DT, Gloer JB: **Phaeofurans and sorbicillin analogues from a fungicolous Phaeoacremonium species (NRRL 32148)**. *Journal of natural products* 2006, **69**(1):113-117.
4. Zhao PJ, Li GH, Shen YM: **New chemical constituents from the endophyte Streptomyces species LR4612 cultivated on Maytenus hookeri**. *Chemistry & biodiversity* 2006, **3**(3):337-342.
5. Cram DJ: **Mold metabolites; the structure of sorbicillin, a pigment produced by the mold Penicillium notatum**. *Journal of the American Chemical Society* 1948, **70**(12):4240-4243.
6. Maskey RP, Grun-Wollny I, Laatsch H: **Sorbicillin analogues and related dimeric compounds from Penicillium notatum**. *Journal of natural products* 2005, **68**(6):865-870.
7. Fahad AA, Abood A, Fisch KM, Osipow A, Davison J, Avramovic M, Butts CP, Piel J, Simpson TJ, Cox RJ: **Oxidative dearomatisation: the key step of sorbicillinoid biosynthesis**. *Chemical science* 2014, **5**(2):523-527.
8. Nicolaou KC, Simonsen KB, Vassilikogiannakis G, Baran PS, Vidali VP, Pitsinos EN, Couladouros EA: **Biomimetic Explorations Towards the Bisorbicillinoids: Total Synthesis of Bisorbicillinol, Bisorbibutenolide, and Trichodimerol**. *Angewandte Chemie* 1999, **38**(23):3555-3559.
9. Abe N, Hirota A: **Chemical studies of the radical scavenging mechanism of bisorbicillinol using the 1,1-diphenyl-2-picrylhydrazyl radical**. *Chemical communications* 2002(6):662-663.
10. Kontani M, Sakagami Y, Marumo S: **First β-1,6-glucan biosynthesis inhibitor, bisvertinolone isolated from fungus, Acremonium strictum and its absolute stereochemistry**. *Tetrahedron Letters* 1994, **35**(16):2577-2580.
11. Bringmann G, Lang G, Gulder TAM, Tsuruta H, Mühlbacher J, Maksimenka K, Steffens S, Schaumann K, Stöhr R, Wiese J *et al*: **The first sorbicillinoid alkaloids, the antileukemic sorbicillactones A and B, from a sponge-derived Penicillium chrysogenum strain**. *Tetrahedron* 2005, **61**(30):7252-7265.
12. Bringmann G, Lang G, Muhlbacher J, Schaumann K, Steffens S, Rytik PG, Hentschel U, Morschhauser J, Muller WE: **Sorbicillactone A: a structurally unprecedented bioactive novel-type alkaloid from a sponge-derived fungus**. *Progress in molecular and subcellular biology* 2003, **37**:231-253.
13. Volp KA, Johnson DM, Harned AM: **A concise synthetic approach to the sorbicillactones: total synthesis of sorbicillactone A and 9-epi-sorbicillactone A**. *Organic letters* 2011, **13**(17):4486-4489.
14. Avramović M: **Analysis of the Genetic Potential of the Sponge-derived Fungus Penicillium Chrysogenum E01 - 10/3 for Polyketide Production. (Doctoral dissertation, Rheinischen Friedrich-Wilhelms-Universität Bonn). Retrieved from http://hss.ulb.uni-bonn.de/2011/2454/2454.htm**; 2011.
15. Bringmann G, Gulder TA, Lang G, Schmitt S, Stohr R, Wiese J, Nagel K, Imhoff JF: **Large-scale biotechnological production of the antileukemic marine natural product sorbicillactone A**. *Marine drugs* 2007, **5**(2):23-30.

16. Weber SS, Kovalchuk A, Bovenberg RA, Driessen AJ: **The ABC transporter ABC40 encodes a phenylacetic acid export system in Penicillium chrysogenum**. *Fungal genetics and biology : FG & B* 2012, **49**(11):915-921.
17. Kovalchuk A, Weber SS, Nijland JG, Bovenberg RA, Driessen AJ: **Fungal ABC transporter deletion and localization analysis**. *Methods in molecular biology* 2012, **835**:1-16.
18. van den Berg MA, Albang R, Albermann K, Badger JH, Daran JM, Driessen AJ, Garcia-Estrada C, Fedorova ND, Harris DM, Heijne WH *et al*: **Genome sequencing and analysis of the filamentous fungus Penicillium chrysogenum**. *Nature biotechnology* 2008, **26**(10):1161-1168.
19. Harris DM, van der Krogt ZA, Klaassen P, Raamsdonk LM, Hage S, van den Berg MA, Bovenberg RA, Pronk JT, Daran JM: **Exploring and dissecting genome-wide gene expression responses of Penicillium chrysogenum to phenylacetic acid consumption and penicillinG production**. *BMC genomics* 2009, **10**:75.
20. Abdel-Lateff A: **Bioproduction of sorbicillin derivatives from marine Trichoderma sp**. *Zeitschrift fur Naturforschung C, Journal of biosciences* 2008, **63**(9-10):631-635.
21. Guo W, Peng J, Zhu T, Gu Q, Keyzers RA, Li D: **Sorbicillamines A-E, nitrogen-containing sorbicillinoids from the deep-sea-derived fungus Penicillium sp. F23-2**. *Journal of natural products* 2013, **76**(11):2106-2112.
22. Peng J, Zhang X, Du L, Wang W, Zhu T, Gu Q, Li D: **Sorbicatechols A and B, antiviral sorbicillinoids from the marine-derived fungus Penicillium chrysogenum PJX-17**. *Journal of natural products* 2014, **77**(2):424-428.
23. Washida K, Abe N, Sugiyama Y, Hirota A: **Novel secondary metabolites, spirosorbicillinols a, B, and C, from a fungus**. *Bioscience, biotechnology, and biochemistry* 2009, **73**(6):1355-1361.
24. Andrade R, Ayer WA, Mebe PP: **The metabolites of Trichoderma longibrachiatum. Part1. Isolation of the metabolites and the structure of trichodimerol**. *Canadian Journal of Chemistry* 1992, **70**(10):2526-2535.